BEBR

Three Alternative Errors-in-Variable Estimation
Methods: A Review and Extension

Cheng-few Lee

BEBR

Three Alternative Errors-in-Variable
Estimation Methods: A Review and Extension

Cheng-few Lee, Professor
Department of Finance

Abstract

Three alternative errors-in-variable (EV) estimation methods, i.e., classical EV method, the mathematical programming EV method and the grouping EV method are reviewed and extended to multiple regression cases.

I.   Introduction

In discussing the statistical law of demand, Schultz (1925)
pointed out that the traditional ordinary least squares (OLS) can
not be used to find a unique linear demand curve when both
explained and explanatory variables are measured with errors.
Frisch (1943), Koopmans (1937) and Allen (1939) have demonstrated
that a priori information is required to determine the true regres-
sion line in an errors-in-variables model. Although the measure-
ment error problem has been clearly pointed out by economists,
econometricians have spent most of their efforts on the specifica-
tion analysis instead of the problems associated with errors of
measurement.

For the single equation model, Wald (1940) has developed a
bi-variate grouping method to determine the straight line without
requiring a priori information. In addition, Reiersol (1941) has
suggested to take the lagged value as an instrumental variable
when the true component instead of the measurement error component
of the observation is serially correlated; Tintner (1952) has
developed the variate difference method for the errors-in-variables
model; Durbin (1954) has shown that the rank order of the explana-
tory variable can be employed as a useful instrumental variable.

For the simultaneous equation model, the main contributions
are due to Anderson and Hurwicz (1949), Konijn (1962), Zellner

(1970) and Goldberger (1972). They combined the a priori infor-
mation required to identify the structural parameters and the
a priori information required to identify the measurement error
variance to deal with problems associated with measurement errors.

In the last three decades, most of the investigation of the
errors-in-variables model has assumed knowledge of the measurement
error variance. In the latest investigations, attempts have been
made to estimate the error variance itself. Needless to say, the
relevance of these empirical results should be carefully evaluated
before drawing misleading conclusions about economic theory.

In this paper, first the effects of measurement error on the
multiple regression coefficients will be investigated in detail to
show that (i) the quality of economic data should be improved to
increase the usefulness of econometric research and (ii) the
standard estimation method should be modified to identify the
estimate of multiple regression coefficients. Secondly, the
classical method, the mathematical programming method and the
grouping method[1] will be extended from the original simple regres-
sion case to the multiple regression case. In the second section,
the basic multiple regression model will be defined and the effects

---

[1] The mathematical programming method refers the method using
iteration procedure to estimate the parameters. These three
different methods will be self explained in sections III, IV and
V respectively.

of measurement error will be investigated in detail. In the third section, the bi-variante classical method will be extended to the multivariate case. Both the mathematical programming method and the grouping method will be extended from the vi-variante case to the multivariate in section IV and V respectively. Finally, the results will be compared and the possible further research in this direction will be indicated.

II. The Effects of Measurement Error on the Multiple Regression Coefficient.

In this section, first a basic model for multiple regression will be defined. Secondly, the effects of measurement error in the multiple regression coefficient will be investigated in detail. Following Moran (1971), a trivariate structural relationship can be specified as

$$W_i = \alpha + \beta U_i + \gamma V_i, \tag{1}$$

where W_i, U_i and V_i are unobserved, but we can observe $Z = W_i + \tau_i$, $X_i = U_i + \varepsilon_i$ and $Y_i = V_i + \eta_i$. U_i and V_i have a joint normal distribution with variances of σ_U^2 and σ_V^2 and correlation coefficient ρ_{UV}. In the observed variables X, Y and Z, the observed errors ε, η and τ are independent normal variables with zero means and variance σ_1^2, σ_2^2 and σ_3^2.

In addition, we also assume that X, Y and Z have a multivariate normal distribution with parameters as follows:

(a) $m_1 = E(X)$

(b) $m_2 = E(Y)$

(c) $m_3 = \alpha + \beta m_1 + \gamma m_2$

(d) $m_{XX} = \sigma_U^2 + \sigma_1^2$

(e) $m_{YY} = \sigma_V^2 + \sigma_2^2$ (2)

(f) $m_{ZZ} = \beta^2 \sigma_U^2 + \gamma^2 \sigma_V^2 + 2\beta\gamma\rho\sigma_U\sigma_V + \sigma_3^2$

(g) $m_{XY} = \rho\sigma_U\sigma_V$

(h) $m_{XZ} = \beta\sigma_U^2 + \gamma\rho\sigma_U\sigma_V$

(i) $m_{YZ} = \beta\rho\sigma_U\sigma_V + \gamma\sigma_V^2.$

The joint sufficient statistics of m_1, m_2, m_3, m_{XX}, m_{YY}, m_{ZZ}, m_{XY}, m_{XZ}, m_{YZ} can be defined as:

(a) $\bar{X} = n^{-1} \Sigma X_i$

(b) $\bar{Y} = n^{-1} \Sigma Y_i$

(c) $\bar{Z} = n^{-1} \Sigma Z_i$

(d) $S_{XX} = n^{-1} \Sigma (X_i - \bar{X})$

(e) $S_{YY} = n^{-1} \Sigma (Y_i - \bar{Y})$ (3)

(f) $S_{ZZ} = n^{-1} \Sigma (Z_i - \bar{Z})$

(g) $S_{XY} = n^{-1} \Sigma (X_i - \bar{X})(Y_i - \bar{Y})$

(h) $S_{XZ} = n^{-1} \Sigma (X_i - \bar{X})(Z_i - \bar{Z})$

(i) $S_{YZ} = n^{-1} \Sigma (Y_i - \bar{Y})(Z_i - \bar{Z}).$

It is well known that the measurement errors will bias the estimate of multiple regression coefficient, but the results have not been carefully investigated. Here, first the effects of

measurement error on the multiple regression coefficients of classical errors-in-variables model will be investigated. Secondly, we will relax the classical assumptions by allowing constant measurement errors, dependence among the true values and measurement errors and dependence among measurement errors. We will call this case the "new classical case" to distinguish it from the "classical case." To identify the regression parameters, without the information of the variance of measurement error, we also try to impose some constraints on the regression coefficients. This will be called the "constrained classical case." The results obtained in this section will be used to explain why both mathematical programming method and the grouping method should be extended to a multiple regression analysis. The effects of measurement error on multiple regression coefficients will be investigated in accordance with the classical case, new classical case and constrained classical case respectively.

From (2), we know that $\bar{X}$, $\bar{Y}$, $\bar{Z}$, S_{XX}, S_{YY}, S_{ZZ}, S_{XY}, S_{XZ} and S_{YZ} are joint sufficient statistics of m_1, m_2, m_3, m_{XX}, m_{YY}, m_{ZZ}, m_{XY}, m_{XZ} and m_{YZ}. If the former nine variables are jointly independent, a set of maximum likelihood equations can be formulated as follows.

(a) $S_{XX} = \sigma_U^2 + \sigma_1^2$

(b) $S_{YY} = \sigma_V^2 + \sigma_2^2$

(c) $S_{ZZ} = \hat{\beta}^2\sigma_U^2 + \hat{\gamma}^2\sigma_V^2 + 2\hat{\beta}\hat{\gamma}\sigma_{UV} + \sigma_3^2$ (4)

(d) $S_{XY} = \sigma_{UV}$

(e) $S_{XZ} = \hat{\beta}\sigma_U^2 + \hat{\gamma}\sigma_{UV}$

(f) $S_{YZ} = \hat{\beta}\sigma_{UV} + \hat{\gamma}\sigma_V^2.$

(1) The Classical Case

From (4), the effects of measurement errors on the estimates of β and γ in the classical case can be seen from the following:

$$\text{plim } \hat{\beta} = \frac{(\sigma_V^2 + \sigma_2^2)\sigma_{WU} - (\sigma_{WV}\sigma_{UV})}{(\sigma_U^2 + \sigma_1^2)(\sigma_V^2 + \sigma_2^2) - (\sigma_{WV})^2} \tag{5}$$

$$\text{plim } \hat{\gamma} = \frac{(\sigma_U^2 + \sigma_1^2)\sigma_{WV} - (\sigma_{WU}\sigma_{UV})}{(\sigma_U^2 + \sigma_1^2)(\sigma_V^2 + \sigma_2^2) - (\sigma_{UV})^2}. \tag{6}$$

From both (5) and (6), the asymtotic biases of $\hat{\beta}$ and $\hat{\gamma}$ can be defined as:

$$\text{plim } \hat{\beta}-\beta = \frac{\sigma_{WU}\sigma_2^2 - \beta(\sigma_U^2\sigma_2^2 + \sigma_V^2\sigma_1^2 + \sigma_1^2\sigma_2^2)}{(\sigma_U^2\sigma_V^2 - \sigma_V^2) + \sigma_U^2\sigma_2^2 + \sigma_V^2\sigma_1^2 + \sigma_1^2\sigma_2^2} \tag{7}$$

$$\text{plim } \hat{\gamma}-\gamma = \frac{\sigma_{WV}\sigma_1^2 - \gamma(\sigma_U^2\sigma_2^2 + \sigma_V^2\sigma_1^2 + \sigma_1^2\sigma_2^2)}{(\sigma_U^2\sigma_V^2 - \sigma_{UV}^2) + \sigma_U^2\sigma_2^2 + \sigma_V^2\sigma_1^2 + \sigma_1^2\sigma_2^2}. \tag{8}$$

The direction of the biases of $\hat{\beta}$ and $\hat{\gamma}$ can be treated according to the following:

Under the assumption that Cov (UV) = 0

i) $\sigma_1^2 = 0$, $\sigma_2^2 > 0$

$$\text{(a) plim } \hat{\beta}-\beta = \frac{\sigma_2^2(\sigma_{WU} - \beta\sigma_U^2)}{\sigma_U^2(\sigma_V^2 + \sigma_2^2)} = 0$$

(b) $\text{plim } \hat{\gamma} - \gamma = \dfrac{-\gamma \sigma_U^2 \sigma_2^2}{\sigma_U^2 (\sigma_V^2 + \sigma_2^2)}$

$$= \dfrac{-\gamma \sigma_2^2}{(\sigma_V^2 + \sigma_2^2)} \qquad \qquad (9)$$

(9a) implies that $\hat{\beta}$ is an asymtotic unbiased estimator of β; (9b) implies that $\hat{\gamma}$ is downward biased estimator of γ.

ii)   If $\sigma_1^2 > 0$, $\sigma_2^2 = 0$

(a) $\text{plim } \hat{\beta} - \beta = \dfrac{-\beta \sigma_1^2}{\sigma_V^2 (\sigma_U^2 + \sigma_1^2)}$

(b) $\text{plim } \hat{\gamma} - \gamma = \dfrac{\sigma_1^2 (\sigma_{WV} - \gamma \sigma_V^2)}{\sigma_V^2 (\sigma_U^2 + \sigma_1^2)} = 0. \qquad (10)$

In accordance with (9), (10) can be used to draw some conclusions about the biases of both $\hat{\beta}$ and $\hat{\gamma}$.

iii) Finally, if $\sigma_1^2 \gtrless 0$, $\sigma_2^2 > 0$

(a) $\text{plim } \hat{\beta} - \beta = - \dfrac{\beta \sigma_1^2}{\sigma_U^2 + \sigma_1^2}$

(b) $\text{plim } \hat{\gamma} - \gamma = - \dfrac{\gamma \sigma_2^2}{\sigma_V^2 + \sigma_2^2}. \qquad (11)$

In this case, the biases of $\hat{\beta}$ and $\hat{\gamma}$ are identical to those biases of the coefficients, which are obtained from regressing Z on X and Z on Y respectively.

Suppose now that Cov (UV) $\neq$ 0

i) $\sigma_1^2 = 0$, $\sigma_2^2 > 0$

   (a)   plim $\hat{\beta} - \beta = \gamma b_{VU} \dfrac{\sigma_2^2}{\sigma_2^2 + \sigma_V^2(1 - R_{UV}^2)}$

   (b)   plim $\hat{\gamma} - \gamma = \dfrac{-\gamma \sigma_2^2}{\sigma_V^2 - b_{VU} + \sigma_2^2}$,   (12)

where b_{VU} is the auxiliary regression coefficient of a regressing
V on U, and R_{UV}^2 is the correlation coefficient between U and V.
(12a) implies that the direction of the bias of $\hat{\beta}$ depends upon the
sign of both γ and b_{VU}; (12b) implies that $\hat{\gamma}$ is a downward biased
estimator of γ unless $(\sigma_V^2 - b_{VU} + \sigma_2^2)$ is smaller than zero.

ii)   If $\sigma_1^2 > 0$, $\sigma_2^2 = 0$

   (a)   plim $\hat{\beta} - \beta = \dfrac{-\beta \sigma_1^2}{\sigma_U^2 - b_{UV} + \sigma_1^2}$

   (b)   plim $\hat{\gamma} - \gamma = b_{UV} \dfrac{\sigma_1^2}{\sigma_1^2 + \sigma_U^2(1 - R_{UV}^2)}$ (13)

where b_{UV} = the auxiliary regression coefficient of regressing U
on V.

iii)   If $\sigma_1^2 > 0$, $\sigma_2^2 > 0$

$$(a) \quad \text{plim } \hat{\beta} - \beta = \frac{\gamma b_{VU} - \dfrac{\beta}{\sigma_U^2}(\sigma_V^2\sigma_1^2 + \sigma_1^2\sigma_2^2)}{\sigma_V^2 - b_{VU} + \sigma_2^2 + \dfrac{\sigma_V^2\sigma_1^2 + \sigma_1^2\sigma_2^2}{\sigma_U^2}}$$

$$(b) \quad \text{plim } \hat{\gamma} - \gamma = \frac{\beta b_{UV} - \dfrac{\gamma}{\sigma_V^2}(\sigma_U^2\sigma_2^2 + \sigma_1^2\sigma_2^2)}{\sigma_U^2 - b_{UV} + \sigma_1^2 + \dfrac{\sigma_V^2\sigma_1^2 + \sigma_1^2\sigma_2^2}{\sigma_V^2}}. \tag{14}$$

From (14), we can see that the direction of the biases of both $\hat{\beta}$ and $\hat{\gamma}$ are ambiguous.

An example may be useful at this point. Suppose that we wish to determine whether a real term or a nominal term should be employed to estimate the parameters of the money demand function when only the deflator is subject to measurement error. Suppose we have a money demand function defined as

$$\frac{M_i}{P_i} = \left(\frac{Y_i}{P_i}\right)^\alpha e^{W_i} \tag{15}$$

where M_i = money demand,

P_i = wholesale price index,

α = constant,

$w_i \sim (0, \sigma^2)$,

P_i is unobserved, but we can observe $P_i' = P_i\varepsilon_i$. By logarithm transformation, an observable money demand can be written as

10

$$\log M_i = \alpha \log Y_i + (1-\alpha)[\log P_i + \log \varepsilon_i] + w_i \qquad (16)$$

If (16) follows the assumptions of classical errors-in-variables model, then a criteria for determining whether the real term or the norminal term is to be employed in this kind of money demand function can be derived as follows:

i) if $\log P_i$ is ommitted, the specification bias of the estimate of α can be written as:

$$\text{plim } \hat{\alpha} - \alpha = (1-\alpha) \frac{\text{cov}(\log P, \log Y)}{\text{var}(\log Y)} \qquad (17)$$

ii) if $\log P_i$ is included, the asymptotic bias of the estimate of α is

$$\text{plim } \hat{\alpha} - \alpha = \frac{(1-\alpha) \text{ var}(\log \varepsilon_i)}{\text{var}(\log Y) + \text{var}(\log \varepsilon_i)} \qquad (18)$$

From (17) and (18), it is clear that the real term instead of the nominal term should be employed to estimate the parameter of money demand function if

$$\left| \frac{\text{cov}(\log P, \log Y)}{\text{var}(\log Y)} \right| > \frac{\text{var}(\log \varepsilon)}{\text{var}(\log Y) + \text{var}(\log \varepsilon)} \qquad (19)$$

Similarly, for a multiple regression, the criteria for either including or excluding price index in the regression can also be drawn by specification and errors-in-variable analysis.

(2) The New Classical Case

We will investigate the effects of measurement error on regression coefficients for both simple and multiple regressions

when the classical assumptions are dropped. For simple regression, we will also derive the effects of constant measurement error on the estimate of the intercept in addition to investigating the effects of random measurement error on the estimate of the slope; for multiple regression, only the effects of random measurement error on the regression coefficients will be investigated.

For bi-variate normal case, if the classical assumptions are dropped, then following (1), we have:

(a) $m_1 = E(X) = E(U) + \bar{\varepsilon}$

(b) $m_2 = E(Y) = \alpha + \beta E(U) + \bar{\eta}$

(c) $m_{XX} = var(X) = \sigma^2 + \sigma_1^2 + 2\sigma_{U\varepsilon}$

(d) $m_{YY} = var(Y) = \beta^2\sigma^2 + \sigma_2^2 + 2\beta\sigma_{U\eta}$

(e) $m_{XY} = cov(XY) = \beta\sigma^2 + \sigma_{V\varepsilon} + \sigma_{U\eta} + \sigma_\varepsilon$ (20)

where

$\bar{\varepsilon}$ is the constant measurement error of U,

$\bar{\eta}$ is the constant measurement error of V,

$\sigma_{U\varepsilon}$, $\sigma_{U\eta}$, $\sigma_{V\varepsilon}$ and $\sigma_{\varepsilon\eta}$ are covariance for relevant variables.

From (20), we can obtain the asymptotic biases for both β and α as

(a) $\text{plim } \hat{\beta}-\beta = \dfrac{\sigma_{V\varepsilon} + \sigma_{U\eta} + \sigma_{\varepsilon\eta} - 2\beta\sigma_{U\varepsilon} - \beta\sigma_1^2}{\sigma^2 + 2\sigma_{U\varepsilon} + \sigma_1^2} = \phi$

(b) $\text{plim } \hat{\alpha}-\alpha = (\bar{\varepsilon} - \beta\bar{\eta}) - \phi[E(V) + \bar{\eta}]$. (21)

The differences between the asymptotic biases of the classical case and the new classical case have following differences.

(i)   In the classical case, the estimate of the slope will be biased toward zero only when the independent variable is measured with error; in the new classical case, having either independent or dependent variables measured with error is a sufficient condition to bias the estimate of the slope.

(ii) Under the classical case, measurement errors will never change the sign of the slope, but in the new classical case, the sign of slope can be changed due to measurement errors, i.e., $\text{plim} \frac{(\hat{\beta}-\beta)}{\beta}$ does not necessarily lie between -1 and 0.

(iii) The intercept has a higher probability of being affected by measurement errors in the new classical case than in the classical case.

For multivariate normal case, (5) and (6) can be rewritten as:

$$\text{plim } \hat{\beta} = \frac{(\sigma_V^2 + \sigma_2^2)\sigma_{WU}' - (\sigma_{WV}'\sigma_{UV}')}{(\sigma_U^2 + \sigma_1^2)(\sigma_V^2 + \sigma_2^2) - (\sigma_{UV}')^2} \tag{22}$$

$$\text{plim } \hat{\gamma} = \frac{(\sigma_U^2 + \sigma_1^2)\sigma_{WV}' - (\sigma_{WU}'\sigma_{UV}')}{(\sigma_U^2 + \sigma_1^2)(\sigma_V^2 + \sigma_2^2) - (\sigma_{UV}')^2}, \tag{23}$$

where $\sigma_{WU}' = \sigma_{WU} + \sigma_{\tau\varepsilon} + \sigma_{W\varepsilon} + \sigma_{\tau U}$,

$\sigma_{WV}' = \sigma_{WV} + \sigma_{\tau\eta} + \sigma_{W\eta} + \sigma_{\tau V}$,

$\sigma_{UV}' = \sigma_{UV} + \sigma_{U\varepsilon} + \sigma_{V\eta} + \sigma_{\varepsilon\eta}$.

If only V is measured with error, then the asymptotic bias of $\hat{\beta}$ can be written as

$$
\text{plim } \hat{\beta} - \beta = \frac{\sigma_{WU}\sigma_2^2 - (\sigma_U^2\sigma_2^2 - \sigma_{U\eta})}{(\sigma_U^2\sigma_V^2 - \sigma_{UV}^2) + \sigma_U^2\sigma_2^2 - \sigma_{U\eta} + 2\sigma_U^2\sigma_{V\eta}}
$$

$$
= \frac{\sigma_2^2[\gamma b_{VU} + \beta R_{U\eta}^2 - 2\beta b_{\eta V}]}{\sigma_2^2(1 - R_U^2 + 2b_{\eta V}) + \sigma_V^2(1 - R_{UV}^2)} . \tag{24}
$$

Under this circumstance, if the variable measured with error is dropped from the regression analysis, the specification bias of β is γb_{VU}. The relative magnitude between γb_{VU} and (plim $\hat{\beta} - \beta$) can be analyzed as follows:

(i)   specifically, if $R_{U\eta}^2 = 0$ and $b_{\eta V} = 0$,
   then (plim $\hat{\beta} - \beta$) < γb_{VU}, $\tag{25}$

(ii) in general, (plim $\hat{\beta} - \beta$) < γb_{VV} if and only if

$$
\frac{1 + \dfrac{\beta R_{U\eta}^2}{\gamma b_{VU}} - \dfrac{2\beta b_{\eta V}}{\gamma b_{VU}}}{(1 - R_{U\eta}^2 + 2b_{\eta U}) + \dfrac{\sigma_V^2(1 - R_{UV}^2)}{\sigma_2^2}} < 1. \tag{26}
$$

Incidentally, both McCallum (1972) and Wickens (1972) have drawn some conclusions about the trade off between the asymptotic bias and specification bias of $\hat{\beta}$. It can be noted that their conclusions are only a special case of (26).

(3)   Constrained Classical Case

If we impose $\beta + \gamma = 1$, on the regression coefficients of (4) and for simplicity, assume only that V is measured with error, then (12a) should be rewritten as

$$\text{plim } \hat{\beta}-\beta = \frac{\sigma_\eta^2(1 - \beta)}{\sigma_U^2 + \sigma_V^2 - 2\sigma_{UV} + \sigma_\eta^2}. \tag{27}$$

The asymptotic bias of this case can be analyzed as follows:

(i)   If $\beta > 0$, then $\hat{\beta}$ is unbiased when $\beta = 1$, $\hat{\beta}$ is upward biased when $\beta < 1$ and $\hat{\beta}$ is downward biased when $\beta > 1$.

(ii) If $\beta < 0$, then $\hat{\beta}$ is always upward biased.

III.   The Extension of Classical Estimation Method to a Multiple Regression Analysis

(1)   Classical Method

From (4), it is clear that if $\rho = 0$, then U is orthogonal to V.   It is well-known that this multiple regression reduces to two simple regression relationship.   If $\rho \neq 0$ then to identify β and γ, we need to know either the actual values of σ_1^2, σ_2^2 and σ_3^2 or the relative ratios among σ_1^2, σ_2^2 and σ_3^2.   We will investigate the following cases:

(i) $\sigma_2^2 = 0$, $\sigma_1^2 > 0$, $\sigma_3^2 > 0$, and $\sigma_1^2 = \lambda\sigma_3^2$

(ii) $\sigma_1^2 > 0$, $\sigma_2^2 > 0$ and $\sigma_3^2 > 0$

a) σ_1^2 and σ_2^2 are known

b) $\sigma_2^2 = \lambda_1 \sigma_1^2$ and $\sigma_3^2 = \lambda_2 \sigma_1^2$.

(Case i)   Only Z and X are measured with errors.

Since $\sigma_2^2 = 0$, from (4 c), (4 e) and (4 f) we have

$$\hat{\gamma} = \frac{(S_{YZ} - \hat{\beta} S_{XY})}{S_{YY}} .\qquad (28)$$

Substituting (28) into (4 e), give us

$$\sigma_U^2 = \frac{S_{YZ}}{\hat{\beta}} - \frac{S_{XY} S_{YZ} - \hat{\beta}(S_{XY})^2}{\hat{\beta} S_{YY}} \qquad (29)$$

From $\lambda = \dfrac{\sigma_1^2}{\sigma_3^2}$, we also have

$$\sigma_3^2 = \frac{1}{\lambda}(S_{XX} - \sigma_U^2). \qquad (30)$$

Substituting (29), (30) and (4 c) and (4 d) into (4 a) and rearranging

$$-\hat{\beta}^2 \lambda [S_{XZ} - \frac{S_{XY} S_{YZ}}{S_{YY}}] + [-(S_{XX} - \frac{S_{XY}^2}{S_{YY}}) + \lambda(S_{ZZ} - \frac{S_{YZ}^2}{S_{YY}})]\hat{\beta} + (S_{XZ} - \frac{S_{XY} S_{YZ}}{S_{YY}}) = 0.$$

$$(31)$$

The solution of (31) can be derived as follows:

Let

$$-\lambda (S_{XZ} - \frac{S_{XY}S_{YZ}}{S_{YY}}) = k_2 = -\lambda k_o$$

$$- (S_{XX} - \frac{S_{XY}}{S_{YY}}) + \lambda (S_{ZZ} - \frac{S_{YZ}^2}{S_{YY}}) = k_1$$

$$(S_{XZ} - \frac{S_{XY}S_{YZ}}{S_{YY}}) = k_o .$$

Then (31) can be written as

$$k_2 \hat{\beta}^2 + k_1 \hat{\beta} + k_o = 0 \tag{32}$$

There are three cases to consider:

(a) $\lambda \to 0$ when $\sigma_1^2 \to 0$, then in this case from (28), we know that

$$\hat{\beta} = \frac{-S_{XY}S_{YZ} + S_{XZ}S_{YY}}{S_{XX}S_{YY} - S_{XY}^2}, \tag{33}$$

(b) $\lambda \to \infty$ when $\sigma_3^2 \to 0$, in this case from (31), we know that

$$\hat{\beta} = \frac{S_{ZZ}S_{YY} - S_{YZ}^2}{S_{XZ}S_{YY} - S_{XY}S_{YZ}} . \tag{34}$$

(33) and (34) correspond to those of simple regression case.

(c) When both $\sigma_1^2 > 0$ and $\sigma_3^2 > 0$, then

$$\hat{\beta}_{1,2} = \frac{k_1 \pm \sqrt{k_1^2 + 4\lambda k_0^2}}{2\lambda k_0}. \tag{35}$$

Whether $\hat{\beta}_1$ or $\hat{\beta}_2$ should be chosen still remains as an open question. When β is determined, γ can be estimated by (28). After both β and γ are estimated, then α can be estimated by

$$\hat{\alpha} = \overline{Z} - \hat{\beta}\overline{X} - \hat{\gamma}\overline{Y}. \tag{36}$$

(Case ii) When Z, X, and Y are all observed with errors

(a) both σ_1^2 and σ_2^2 are known.

By (4 b)-(4 f), we can obtain the two normal equations as follows:

(a´) $S_{XZ} = \beta(S_{XX} - \hat{\sigma}_1^2) + \gamma S_{XY}$

(b´) $S_{YZ} = \beta S_{XZ} + \gamma(S_{YY} - \hat{\sigma}_2^2).$ $\tag{37}$

Solving (37) by Cramer's rule we have

$$\hat{\beta} = \frac{S_{XZ}S_{YY} - S_{XY}S_{YZ} - S_{XZ}\hat{\sigma}_2^2}{S_{XX}S_{YY} - \hat{\sigma}_1^2 S_{YY} - \hat{\sigma}_2^2 S_{YY} + \hat{\sigma}_1^2\hat{\sigma}_2^2 - (S_{XY})^2}$$

$$\hat{\gamma} = \frac{S_{YZ}S_{XX} - S_{XZ}S_{XY} - S_{YZ}\hat{\sigma}_1^2}{S_{XX}S_{YY} - \hat{\sigma}_1^2 S_{YY} - \hat{\sigma}_2^2 S_{XX} + \hat{\sigma}_1^2\hat{\sigma}_2^2 - (S_{XY})^2}. \tag{38}$$

(b) both $\sigma_2^2 = \lambda_1 \sigma_1^2$ and $\sigma_3^2 = \lambda_2 \sigma_1^2$ are known.

From (4 a), (4 e) and (4 f), we have

$$S_{ZZ} = \hat{\beta}S_{XZ} + \hat{\gamma}S_{YZ} + \sigma_3^2. \tag{39}$$

From (4 b) and (4 e) we have

$$\sigma_3^2 = \lambda_2 \sigma_1^2 = \lambda_2 [S_{XX} - \frac{S_{XZ} - \gamma S_{XY}}{\beta}]. \tag{40}$$

Substituting (40) into (39) and rearranging, we have

$$\hat{\gamma} = \frac{\{\beta[S_{ZZ} - \lambda_2 S_{XX}] - \beta^2 S_{XZ} + \lambda_2 S_{XZ}\}}{S_{YZ} + \lambda_2 S_{XY}} \tag{41}$$

From (4 b), (4 c), (4 e) and (4 f) we will obtain

$$\hat{\beta}^2 S_{XY} - \hat{\beta}S_{YZ} + \hat{\gamma}[\hat{\beta}(S_{YY} - \lambda_1 S_{XX}) + \lambda_1 S_{XZ}] - \hat{\gamma}^2 \lambda_1 S_{XY} = 0. \tag{42}$$

Substituting (41) into (42) and rearranging, we have

$$(\hat{\beta}^2 S_{XY} - \hat{\beta}S_{YZ})(\hat{\beta}S_{YZ} + \lambda_2 S_{XY})^2 + [\hat{\beta}T - \hat{\beta}^2 S_{XZ} + \lambda_2 S_{XZ}][\hat{\beta}M + \lambda_2 S_{XZ}]$$

$$[\hat{\beta}S_{YZ} + \lambda_2 S_{XY}] - \lambda_1 S_{XY}[\hat{\beta}T - \hat{\beta}^2 S_{XZ} + \lambda_2 S_{XZ}]^2 = 0, \tag{43}$$

where $T = S_{ZZ} - \lambda_2 S_{XX}$

$M = S_{YY} - \lambda_1 S_{XX}$.

After rearranging, (43) will become the following cubic equation

$$\hat{\beta}^3 H_3 + \hat{\beta}^2 H_2 + \hat{\beta}H_1 + H_o = 0, \tag{44}$$

where

$$H_3 = S_{XY}S_{YZ}^2 - MS_{XZ}S_{YZ} - \lambda_1 S_{XY}S_{XZ}^2$$

$$H_2 = -S_{YZ}^2 + 2\lambda_2 S_{XY}^2 S_{YZ} + MTS_{YZ} - M\lambda_2 S_{XY}S_{YZ} - \lambda_1 S_{XZ}^2 S_{YZ} + 2\lambda_1 TS_{XZ}S_{XY}$$

$$H_1 = \sigma_2^2 S_{XY}^3 - 2\lambda_2 S_{XY}S_{YZ}^2 + MT\lambda_2 S_{XY} + \lambda_2 S_{XZ}S_{YZ} + \lambda_2 MS_{YZ}S_{XZ}$$
$$\quad - \lambda_1\lambda_2 S_{XZ}^2 S_{XY} + 2\lambda_1 S_{XY}S_{XZ}^2 - \lambda_1 T^2 S_{XY}$$

$$H_0 = \lambda_2^2 S_{YZ}S_{XY}^2 + \lambda_2^2 S_{XZ}S_{XY} + \lambda_1\lambda_2 S_{XZ}^2 S_{YZ} + T\lambda_1\lambda_2 S_{XZ}S_{YZ} - 2\lambda_1\lambda_2 TS_{XZ}S_{XY}.$$

When $\lambda_2 = 0$, (44) will reduce to a quadratic equation in $\hat{\beta}$.

(2) Constrained Classical Method

Under the classical case (Case ii), if we only know $\hat{\sigma}_2 = \hat{\sigma}_1\lambda$, then we can identify β and γ by imposing $\beta + \gamma = 1$.
From (4 b) and (4 c) we have

$$-\sigma_V^2 = -\lambda\sigma_U^2 - S_{YY} + \lambda S_{XX}, \tag{45}$$

also from (4 d), (4 e), (4 f) and (45), we have

$$S_{YZ} = (1 - \beta)(-\lambda\sigma_U^2 - S_{YY} + \lambda S_{XX}) + \beta S_{XY} \tag{46}$$

$$S_{XZ} = \beta\sigma_U^2 + (1 = \beta)S_{XY}. \tag{47}$$

From (46) and (47), we can obtain a quadratic equation in $\hat{\beta}$

$$\hat{\beta}^2\{(1-\lambda)S_{XY} - S_{YY}\} + \beta(-S_{YZ} - \lambda S_{XZ} + 2\lambda S_{XY} + S_{YY} - \lambda S_{XX}) + \lambda(S_{XZ} - S_{XY}) = 0. \tag{48}$$

When $\lambda = 0$, then (48) will reduce to

$$\hat{\beta} S_{XY} = -[S_{YZ} + S_{YY}]$$

i.e. $\hat{\beta} = \dfrac{S_{YZ} - S_{YY}}{S_{XY} - S_{YY}}$ (49)

Imposing $\beta + \gamma = 1$, upon a multiple regression will help to identify the regression coefficients, but it should also be realized that the constrained regression technique will bias the estimates of the regression coefficients if the unrestricted estimator fails to satisfy the restriction $\beta + \gamma = 1$. The advantages and disadvantages of the constrained regression technique has been discussed by Theil (1971) in some detail.

IV. The Mathematical Programming Method

The approach involves the estimation of the parameters of a function conditional on the maximum likelihood function adjusted for the true values. This method is different from the classical method as follows:

(i) The variances of measurement errors for every observation are different.

(ii) A weighted regression method is applied.

(iii) An iteration procedure is used to obtain a consistent estimator.

It can be proved that the mathematical programming method reduces to the classical method under certain conditions.

The principal contribution I make here is to extend the bi-variate mathematical programming method (which was developed by Deming (1943), York (1966) and Clutton-Brock (1967)) to a tri-variate case.

Following York (1966), we define $W(Z_i)$, $W(X_i)$, and $W(Y_i)$ which are the weights of the various observations of Z_i, X_i and Y_i. It is assumed W, U, and V are functionally rather than structurally related.[2] Also, the adjusted values of Z_i, X_i and Y_i, based upon the measurement error variances of every observation and maximum likelihood procedure are defined as

$$x_i = \frac{A_o B_1 + A_1 B_o}{A_1 B_2 - A_2 B_1} \tag{50}$$

$$y_i = \frac{A_2 B_o + A_o B_2}{A_1 B_2 - A_2 B_1} \tag{51}$$

$$z_i = \alpha + \beta x_i + \gamma y_i, \tag{52}$$

$$(i = 1, 2, \ldots, N)$$

where $A_2 = \sigma_{\tau i}^2 + \beta^2 \sigma_{\varepsilon i}^2$

$A_1 = \beta \gamma \sigma_{\varepsilon i}^2$

$A_o = \sigma_{\tau i}^2 X_i + \alpha \beta \sigma_{\eta i}^2 - \beta \sigma_{\varepsilon i}^2 Z_i$

$B_2 = \beta \gamma \sigma_{\eta i}^2$

---

[2] The differences between the functional and the structural relationship can be found in Grabill (1961), Kendall and Stuart (1961) and Moran (1971).

$$B_1 = \sigma^2_{\tau i} + \gamma^2 \sigma^2_{\eta i}$$

$$B_o = \sigma^2_{\tau i} Y_i + \alpha\beta\sigma^2_{\eta i} - \gamma\sigma^2_{\eta i} Z_i.$$

The mathematical programming procedure begins by minimizing

$$S = \Sigma_i \{W(X_i)(x_i - X_i)^2 + W(Y_i)(y_i - Y_i)^2 + W(Z_i)(z_i - Z_i)^2\} \qquad (53)$$

S.t. $\quad z_i = \alpha + \beta x_i + \gamma y_i.$ \hfill (54)

Totally differentiating (53) and (54), we have

$$\delta S = \Sigma_i \{W(X_i)(x_i - X_i)\delta x_i + W(Y_i)(y_i - Y_i)\delta y_i$$

$$+ W(Z_i)(z_i - Z_i)\delta z_i\} \qquad (55)$$

$$-\delta z_i + \delta\alpha + x_i \delta b + \beta\delta x_i + y_i \delta\gamma + \gamma\delta y_i = 0, \qquad (56)$$

where δ is the operator of total differentiation. Let
(55) + λ_i(56) = 0, we now have

$$\Sigma_i\{\delta x_i[W(X_i)(x_i - X_i) + \beta\lambda_i]\} + \Sigma_i\{\delta y_i[W(Y_i)(y_i - Y_i) + \gamma\lambda_i]\}$$

$$+\Sigma_i\{\delta z_i[W(Z_i)(z_i - Z_i) - \lambda_i]\} + \delta\alpha\Sigma_i\lambda_i + \delta\beta\Sigma_i\lambda_i x_i + \delta\gamma\Sigma_i\lambda_i y_i = 0$$

$$(57)$$

Equating the coefficients to zero gives

$$(a) \quad (x_i - X_i) = \frac{-\lambda_i\beta}{W(X_i)}$$

$$(b) \quad (y_i - Y_i) = \frac{-\lambda_i\beta}{W(Y_i)} \qquad (58)$$

$$(c) \quad (z_i - Z_i) = \frac{\lambda_i}{W(Z_i)}.$$

(a) $\Sigma_i \lambda_i \, x_i = 0$

(b) $\Sigma_i \lambda_i Y_i = 0$ (59)

(c) $\Sigma_i \lambda_i = 0.$

Substituting (58) into (54) yields

$$\lambda_i = K_i (\alpha + \beta X_i + \gamma Y_i - Z_i) \qquad (60)$$

where

$$K_i = \frac{W(X_i) W(Y_i) W(Z_i)}{W(X_i) W(Y_i) + \beta^2 W(Z_i) W(Y_i) + \gamma^2 W(Z_i) W(X_i)} \qquad (61)$$

From (59 c) and (60), we have

$$\Sigma_i K_i (\alpha + \beta X_i + \gamma Y_i - Z_i) = 0 \qquad (62)$$

From (59 a), (59 b) and (60), we have

$$\Sigma_i K_i \{\alpha X_i + \beta X_i^2 + \gamma X_i Y_i - X_i Z_i\} - \Sigma_i K_i^2 \{\frac{\beta}{W(X_i)} (\alpha + \beta X_i + \gamma Y_i - Z_i)^2\} = 0.$$

$$(63)$$

$$\Sigma_i K_i \{\alpha Y_i + \beta X_i Y_i + \gamma Y_i^2 - Y_i Z_i\} - \Sigma K_i^2 \{\frac{\gamma}{W(Y_i)} (\alpha + \beta X_i + \gamma Y_i - Z_i)^2\} = 0$$

$$(64)$$

From (62), we have

$$\alpha = \overline{Z} - b\overline{X} - \lambda\overline{Y}, \qquad (65)$$

where $\overline{X} = \dfrac{\Sigma_i X_i K_i}{\Sigma_i K_i}$, $\overline{Y} = \dfrac{\Sigma_i K_i Y_i}{\Sigma_i K_i}$, and $\overline{Z} = \dfrac{\Sigma K_i Z_i}{\Sigma K_i}$.

If we let $P_i = Z_i - \bar{Z}$, $Q_i = X_i - \bar{X}$ and $R_i = Y_i - \bar{Y}$, and substitute them into (60) and (61), we have

$$\Sigma K_i \{\beta Q_i^2 + \gamma Q_i R_i - P_i Q_i\} - \Sigma_i K_i \{\frac{\beta}{W(X_i)} (\beta Q_i + \gamma R_i - P_i)^2\} = 0.$$

$$\text{(66)}$$

$$\Sigma K_i \{\beta Q_i R_i + \gamma R_i^2 - R_i P_i\} - \Sigma K_i^2 \{\frac{\gamma}{W(Y_i)} (\beta Q_i + \gamma R_i - P_i)^2\} = 0.$$

$$\text{(67)}$$

From (66) and (67), two sixth power equations in β and γ can be obtained. These equations have to be solved numerically.

If Y is measured without errors, then (67) reduces to

$$\Sigma K_i \{\beta Q_i R_i + \gamma R_i^2 - P_i R_i\} = 0. \tag{68}$$

From (68) and (66), we can obtain

$$\beta^3 C_3 + 2\beta^2 C_2 + \beta C_1 + C_o = 0, \tag{69}$$

where $C_3 = \Sigma_i \dfrac{K_i^2}{W(X_i)} (Q_i - \dfrac{R_i \Sigma K_i Q_i P_i}{\Sigma K_i R_i^2})^2$

$$C_2 = \Sigma_i \frac{K_i^2}{W(X_i)} (Q_i - R_i \frac{\Sigma K_i Q_i R_i}{\Sigma K_i R_i^2})(P_i - R_i \frac{\Sigma K_i R_i P_i}{\Sigma K_i R_i^2})$$

$$C_1 = \Sigma K_i P_i^2 - \frac{(\Sigma_i K_i Q_i R_i)^2}{\Sigma_i K_i R_i^2} - \Sigma (\frac{K_i^2}{W(X_i)})(\frac{R_i \Sigma K_i R_i P_i}{\Sigma K_i^2} - P_i)^2$$

$$C_o = \frac{\Sigma K_i Q_i R_i \Sigma K_i P_i Q_i}{\Sigma K_i R_i^2} - \Sigma K_i P_i Q_i.$$

When Y_i is orthogonal to X_i, (68) reduces to the "Least Square Cubic" of York (1966). Under some special assumptions, e.g., $\sigma^2_{\tau i} = \sigma^2_\tau$, $\sigma^2_{\varepsilon i} = \sigma^2_\varepsilon$ and $\sigma^2_{\eta i} = \sigma^2_\eta$, the solutions of β in (66) can be shown to be similar to what we have derived in last section.

This extension will reduce to Deming's (1943) weighted regression results when the quadratic term of (66) and (67) are omitted. However, my result is more general than Deming's weighted multiple regression analysis.

V. A Generalization of Wald's Method of Fitting Straight Lines

The relationship between the grouping method and OLS can be demonstrated by applying OLS to either instrumental variable or grouped data.

Gibson and Jowett (1957) and Hooper and Theil (1958) have used the concept of "Centre" (by taking subgroup means) to extend Wald's method of fitting straight lines. However, Gibson and Jowett's method can not be used to treat a model when independent variables are subject to observation errors, since their standard errors for regression coefficients are obtained under the assumption that the independent variables are observed without errors. The Hooper and Theil approach can be shown to reduce to Gibson and Jowett's approach.

Here, a new approach is developed which is based upon specification analysis. This method is useful for econometric research, since it has the advantages of the Gibson-Jowett approach but not its limitations. The main purposes of this sections are: (i) to

derive a stepwise grouping method which can treat n-variate regression problem and (ii) to derive an asymptotic confidence interval for the regression coefficients. In effect, I generalize the Wald grouping method here to the n-variate case.

Gibson and Jowett (1957) have applied Wald's bi-variate grouping method in estimating the multiple regression coefficients when the independent variables are not subject to errors. For simplicity, and without loss generality, we assume that the intercept of (1) is zero, then Gibson and Jowett's method can be described as follows:

By specification analysis, we have

(a) $\quad b = \beta + b_{VU}\gamma$

(b) $\quad r = \gamma + b_{UV}\beta,$ $\hspace{5cm}$ (70)

where $\quad b = (U'U)^{-1}U'W,$

$\quad\quad r = (V'V)^{-1}V'W,$

$\quad\quad b_{VU} = (U'U)^{-1}U'V,$

$\quad\quad b_{UV} = (V'V)^{-1}V'U.$ $\hspace{4cm}$ (71)

From (70), both β and γ can be estimated as

(a) $\quad \hat{\beta} = \dfrac{b - r b_{vu}}{1 - b_{vu} b_{uv}}$

(b) $\quad \hat{\gamma} = \dfrac{r - b b_{uv}}{1 - b_{vu} b_{uv}}$ $\hspace{4cm}$ (72)

Gibson and Jowett employed a bi-variate grouping method instead of OLS to estimate b, r, b_{vu} and b_{uv} to simplify the calculation procedure. If both u and v are measured without error, then both OLS and Wald's grouping method can be used to obtain the consistent estimators for b, r, b_{vu} and b_{uv}; if both u and v are measured with errors, then only Wald's grouping method can be used to the consistent estimators for (71).

Now, we will derive an alternative method when independent variables are either measured with or without errors.

Case (A) - When independent variables are not subject to errors.

In this case, by substituting w = z - τ into (1) we have

$$z_i = \beta u_i + \gamma v_i + \tau_i \tag{73}$$

If we multiply both sides of (71) by an idempotent matrix $M_2 = I - v(v'v)^{-1}v'$, then we have

$$M_2 Z = M_2 u\beta + M_2\tau, \tag{74}$$

where

(a) $M_2 Z = Z - v(v'v)^{-1}v'Z = Z - vr$

(b) $M_2 u\beta = [u - v(v'v)^{-1}v'u]\beta = (u - vb_{uv})\beta \tag{75}$

After estimating r and b_{uv} by bi-variate OLS for (72), we can substitute (75) into (74) to obtain

$$Z_1^* = u^*\beta + \tau_1^*, \tag{76}$$

where

$$Z_1^* = z - v\hat{r},$$

$$u^* = u - v\hat{b}_{uv}, \text{ and}$$

$$\tau_1^* = M_2\tau.$$

From (76), we can obtain the consistent estimate of β by employing the bi-variate OLS.

Similarly, we can obtain

$$Z_2^* = v^*\gamma + \tau_2^*, \tag{77}$$

where $Z_2^* = z - u\hat{b},$

$$v^* = v - u\hat{b}_{vu},$$

$$\tau_2^* = M_1\tau \quad \text{and}$$

$$M_1 = I - u(u'u)^{-1}u'.$$

From (77), we can obtain the consistent estimate of γ by employing the bi-variate OLS.

Case (B) - When independent variables are subject to error.

If both u and v are also measured with errors, then these observation errors will affect the stepwise OLS method in four different (but separable) directions; i.e., they affect (i) the estimate of b and r, (ii) the estimate of b_{uv} and b_{vu}, (iii) the values of u and v individually, and (iv) the estimates of β and γ.

If b, r, b_{uv}, b_{vu}, β and γ are estimated by Wald's bi-variate grouping method instead of OLS method, then the consistency property for these estimators can be preserved. Now, we will prove this argument. For simplicity, we will only deal with the estimation procedure of β. Following (76), a stepwise grouping method (SG) can be employed to replace the stepwise OLS to obtain the consistent estimator of β in accordance with following three steps:

(i) Ordering Y in ascending order and deriving Y and Z into two equal groups, we have

$$\hat{r} = \frac{\bar{Z}_1 - \bar{Z}_2}{\bar{Y}_1 - \bar{Y}_2}$$

$$= \frac{\beta(\bar{u}_1 - \bar{u}_2) + \gamma(\bar{v}_1 - \bar{v}_2) + m^{-1}\left(\sum_{i=1}^{m}\tau_i - \sum_{i=m+1}^{n}\tau_i\right)}{(\bar{v}_1 - \bar{v}_2) + m^{-1}\left(\sum_{i=1}^{m}\eta_i - \sum_{i=m+1}^{n}\eta_i\right)}, \qquad (78)$$

where $m = \frac{1}{2}n$.

Since the variance of the second terms in the numerator and denominator are $0(m^{-1})$, then

$$\plim_{m\to\infty} \hat{r} = \gamma + b_{uv}\beta = r. \qquad (79)$$

(ii)   Ordering Y in ascending order and dividing Y and X into two equal groups, then we can obtain

$$\hat{b}_{uv} = \frac{\overline{X}_1 - \overline{X}_2}{\overline{Y}_1 - \overline{Y}_2}$$

$$= \frac{b_{uv}(\overline{v}_1 - \overline{v}_2) + m^{-1}(\sum_{i=1}^{m} \varepsilon_i - \sum_{i=m+1}^{n} \varepsilon_i)}{(\overline{v}_1 - \overline{v}_2) + m^{-1}(\sum_{i=1}^{m} \eta_i - \sum_{i=m+1}^{n} \eta_i)}. \tag{80}$$

Since the variances of the second terms in the numerator and denominator are $0(m^{-1})$, then

$$\plim_{m \to \infty} \hat{b}_{uv} = b_{uv} \tag{81}$$

Substituting u, v, r and b_{uv} of (81) by X, Y, $\hat{r}$ and $\hat{b}_{uv}$, we have

(a) $Z^* = Z - \hat{r}Y_1 = (W - v\hat{r}) + (\tau - \eta\hat{r}) = W_1^* + \tau_1^*$

(b) $X^* = X - Y\hat{b}_{XY} = (u - v\hat{b}_{XY}) + (\varepsilon - \eta\hat{b}_{XY}) = u^* + \varepsilon^*.$

$$\tag{82}$$

In (72), W* and u* are true values for the observable values Z* and X*. Following (76), the true relationship between W* and u* can be defined as

$$W^* = u^*\beta. \tag{83}$$

To estimate β, we should substitute Z_i^* and X^* for W^* and u^*.

Now, we will examine whether (81), (82) and (83) satisfy Wald's five assumptions.

(a) $\text{plim } (\varepsilon_i^* \varepsilon_j^*) = \text{plim } (\varepsilon_i \varepsilon_j - \varepsilon_i \eta_j \hat{b}_{uv} - \varepsilon_j \eta_i \hat{b}_{uv} + \eta_i \eta_j \hat{b}_{uv}^2) = 0$.

(b) $\text{plim } (\tau_i^* \tau_j^*) = \text{plim } (\tau_i \tau_j - \tau_i \eta_j \hat{r} - \tau_j \eta_i \hat{r} + \eta_i \eta_j \hat{r}^2) = 0$.

(c) $\text{plim } (\varepsilon_i^* \tau_i^*) = \text{plim } (\varepsilon_i \tau_i - \varepsilon_i \eta_i \hat{r} - \eta_i \tau_i + \eta_i^2 \hat{b}_{uv} \hat{r}) = \sigma_\eta^2 b_{uv} r$.

(d) (83) is a single linear relationship.

(e) The limit inferior of

$$\left| \frac{\sum\limits_{i=1}^{m} X_i^* - \sum\limits_{i=m+1}^{n} X_i^*}{m} \right| \qquad (n = 2, 3, \ldots \text{ad. inf.}) \quad (84)$$

is positive.

Since the X_i^*'s (i=1, ..., n) are in ascending order.

(iii) Ordering X^* in ascending order and dividing X^* and Z^* into two equal groups respectively, then we can obtain

$$\hat{\beta} = \frac{\overline{Z}_1^* - \overline{Z}_2^*}{\overline{X}_1^* - \overline{X}_2^*}$$

$$= \frac{\beta(\overline{u}_1^* - \overline{u}_2^*) + m^{-1}\left(\sum\limits_{i=1}^{m} \tau_i^* - \sum\limits_{i=m+1}^{n} \tau_i^* \right)}{(\overline{u}_1^* - \overline{u}_2^*) + m^{-1}\left(\sum\limits_{i=1}^{m} \varepsilon_i^* - \sum\limits_{i=m+1}^{n} \varepsilon_i^* \right)} \qquad (85)$$

Following the arguments in (81) and (79), we can conclude that

$\text{plim } \hat{\beta} = \beta$.

For the two independent variables case, we need to employ a three step bi-variate grouping method to obtain a consistent estimator for each regression coefficient. In general, $\sum_{i=1}^{k} i$ steps are required to estimate each regression coefficient for a k-independent-variables multiple regression.

The estimators of the variance of ε^* and η^* can be derived as follows:

(i) The properties of the relevant statistics, S_{Z*}^2, S_{X*}^2, S_{W*}^2, S_{W*U*} and S_{Z*X*} are

   (a) $\text{plim } (S_{Z*}^2) = \sigma_{W*}^2 + \sigma_{\tau*}^2$

   (b) $\text{plim } (S_{X*}^2) = \sigma_{u*}^2 + \sigma_{\varepsilon*}^2$

   (c) $\text{plim } S_{W*}^2 = \hat{\beta}^2 \sigma_{u*}^2$

   (d) $\text{plim } S_{W*u*} = \hat{\beta} \sigma_u^2$

   (e) $\text{plim } S_{Z*X*} = \sigma_{W*u*}^2 + \sigma_\eta^2 b_{xy}\gamma.$ (86)

(ii) From (86), we can obtain consistent estimators for $\sigma_{\varepsilon*}^2$ and $\sigma_{\tau*}^2$ as

$$\hat{\sigma}_{\varepsilon*}^2 = [S_{X*}^2 - \frac{S_{Z*X*}}{\hat{\beta}} + \frac{\hat{\gamma}\hat{b}_{xy}\hat{\sigma}_\eta^2}{\hat{\beta}}]$$

$$\hat{\sigma}_{\tau*}^2 = [S_{Z*}^2 - \hat{\beta} S_{Z*X*} + \hat{\beta}\hat{\gamma}\hat{b}_{xy}\hat{\sigma}_\eta^2], \qquad (87)$$

where S_{Z*}^2 = sample variance of Z^* observations,

      S_{X*}^2 = sample variance of X^* observations,

      S_{X*Z*} = sample covariance of X^* and Z^* observations.

The Asymptotic Confidence Interval for $\hat{\beta}$ can be derived as follows:

Following Wald (1940), we define

$$(S'_{X*})^2 = \frac{\sum\limits_{i=1}^{m} (X_i^*-\overline{X}_1^*)^2 + \sum\limits_{i=m+1}^{n} (X_i^*-\overline{X}_2^*)^2}{n} \tag{88}$$

$$(S'_{Z*})^2 = \frac{\sum\limits_{i=1}^{m} (Z_i^*-\overline{Z}_1^*)^2 + \sum\limits_{i=m+1}^{n} (Z_i^*-\overline{Z}_2^*)^2}{n}$$

$$S'_{X*Z*} = \frac{\sum\limits_{i=1}^{m} (Z_i^*-\overline{Z}_1^*)(X_i^*-\overline{X}_1^*) + \sum\limits_{i=m+1}^{n} (Z_i^*-\overline{Z}_2^*)(X_i^*-\overline{X}_2^*)}{n},$$

where $\quad \overline{X}_1^* = \dfrac{\sum\limits_{i=1}^{m} X_i^*}{m}$, $\quad \overline{X}_2^* = \dfrac{\sum\limits_{i=m+1}^{n} X_i^*}{m}$

$$\overline{Z}_1^* = \frac{\sum\limits_{i=1}^{m} Z_i^*}{m}, \qquad \overline{Z}_2^* = \frac{\sum\limits_{i=m+1}^{n} Z_i^*}{m}.$$

Under the probability limit sense, following Hogg and Graig (1969), the statistics S'_{X*}, S'_{Z*} and S'_{X*Z*} are distributed independent of $\overline{X}_1^*$, $\overline{X}_2^*$, $\overline{Z}_1^*$ and $\overline{Z}_2^*$.

Now we shall show that

$$\frac{\sqrt{n} \; K \; (\hat{\beta}-\beta)}{\sqrt{(S_{Z*}')^2 + \beta^2 (S_{X*}')^2 - 2\beta S_{Z*X*}'}} \tag{89}$$

is asymtotically normally distributed with zero near and unit variance, where $K = \overline{X}_1^* - \overline{X}_2^*$, then from (79), we have

$$K(\hat{\beta}-\beta) = (\overline{Z}_1^* - \overline{Z}_2^*) - \beta K$$

$$= [(\overline{Z}_1^* - \overline{W}_1^*) - (\overline{Z}_2^* - \overline{W}_2^*)] - \beta(\overline{\epsilon}_1^* - \overline{\epsilon}_2^*)$$

$$= (\overline{\tau}_1^* - \overline{\tau}_2^*) - \beta(\overline{\epsilon}_1^* - \overline{\epsilon}_2^*). \tag{90}$$

If both τ^* and ϵ^* are normally distributed, then from (88) and (90), we can show that

$$\sqrt{n} \; K(\hat{\beta}-\beta) \; \tilde{} \; AN(0, \sigma^2), \tag{91}$$

where

$$\sigma^2 = (\sigma_{\tau*}^2 + \beta^2 \tau_{\epsilon*}^2 - 2\beta b_{uv} r \sigma_\eta^2).$$

Similarly, from (84), (85) and (86), we have

$$(Z_i^* - \beta X_i^*) - (\overline{Z}_1^* - \beta\overline{X}_1^*) = (\tau_i^* - \overline{\tau}_1^*) - \beta(\epsilon_i^* - \overline{\epsilon}_1^*)$$

$$(i=1, \ldots, m)$$

and $(Z_i^* - \beta X_i^*) - (\overline{Z}_2^* - \beta\overline{X}_2^*) = (\tau_i^* - \overline{\tau}_2^*) - \beta(\epsilon_i^* - \overline{\epsilon}_2^*). \tag{92}$

$$(i=m+1, \ldots, n).$$

From (88), we also have

$$\beta (S_{X*}')^2 + (S_{Z*}')^2 - 2S_{X*Z*}'$$

$$= \frac{n}{n-2} \left\{ \frac{\sum\limits_{i=1}^{m} [(Z_i^* - \beta X_i^*) - (\bar{Z}_1^* - \beta \bar{X}_1^*)]^2 + \sum\limits_{i=m+1}^{n} [(Z_i^* - \beta X_i^*) - (\bar{Z}_2^* - \beta \bar{X}_2^*)]^2}{n} \right\} \tag{93}$$

From (92) and (93), we have

$$\text{plim} \{ (S_{Z*}')^2 + \beta^2 (S_{X*}')^2 - 2\beta S_{X*Z*}' \}$$

$$= \sigma_{\tau*}^2 + \beta^2 \sigma_{\varepsilon*}^2 - 2\beta b_{uv} r \sigma_\eta^2 \tag{94}$$

Following (91) and (94), we have

$$\frac{\sqrt{n}\ K(\hat{\beta} - \beta)}{\sqrt{\beta^2 S_{X*}'^2 + S_{Z*}'^2 - 2\beta S_{Z*X*}'}} \sim AN(0,\ 1) \tag{95}$$

Denote by z_0 the critical value of standarized normal distribution to a chosen probability level, then the deviation of $\hat{\beta}$ from an assured population value β is significant if

$$\left| \frac{K(\hat{\beta} - \beta)\ \sqrt{n}}{\sqrt{S_{Z*}'^2 + \beta^2 S_{X*}'^2 - 2\beta S_{Z*X*}'}} \right| \geq z_0 \tag{96}$$

(96) implies that

$$A_2 \beta^2 - 2A_1 \beta + A_0 = 0, \tag{97}$$

where

$$A_2 = \frac{z_o^2}{n} S_{X*}'^2 - K^2$$

$$A_1 = \hat{\beta}K^2 - \frac{z_o}{n} S_{Z*X*}'$$

$$A_o = \frac{z_o^2}{n} S_{Z*}' - K^2\hat{\beta}^2$$

Finally, the standard error of $\hat{\beta}$ can be obtained by

$$\frac{\beta_i^* - \hat{\beta}}{z_o} \qquad (i=1,\ 2) \qquad\qquad (98)$$

where β_i^* = the root obtained from (97).

V. Conclusion

The classical method is a special case of mathematical pro-gramming methods; the grouping method is a special case of the instrumental variable approach; the relationship between the grouping method and the classical method also can be found by applying ordinary least squares to grouping data. Both classical and mathematical programming methods can indicate the magnitude and direction of measurement errors affecting the estimates of regression parameters. However, the grouping method is more applicable at least for econometricians, since it requires much less information; specifically, it does not require knowledge of

the variances, a priori. That is why it is suggested that further works in the errors-in-variables model concentrate on grouping method.
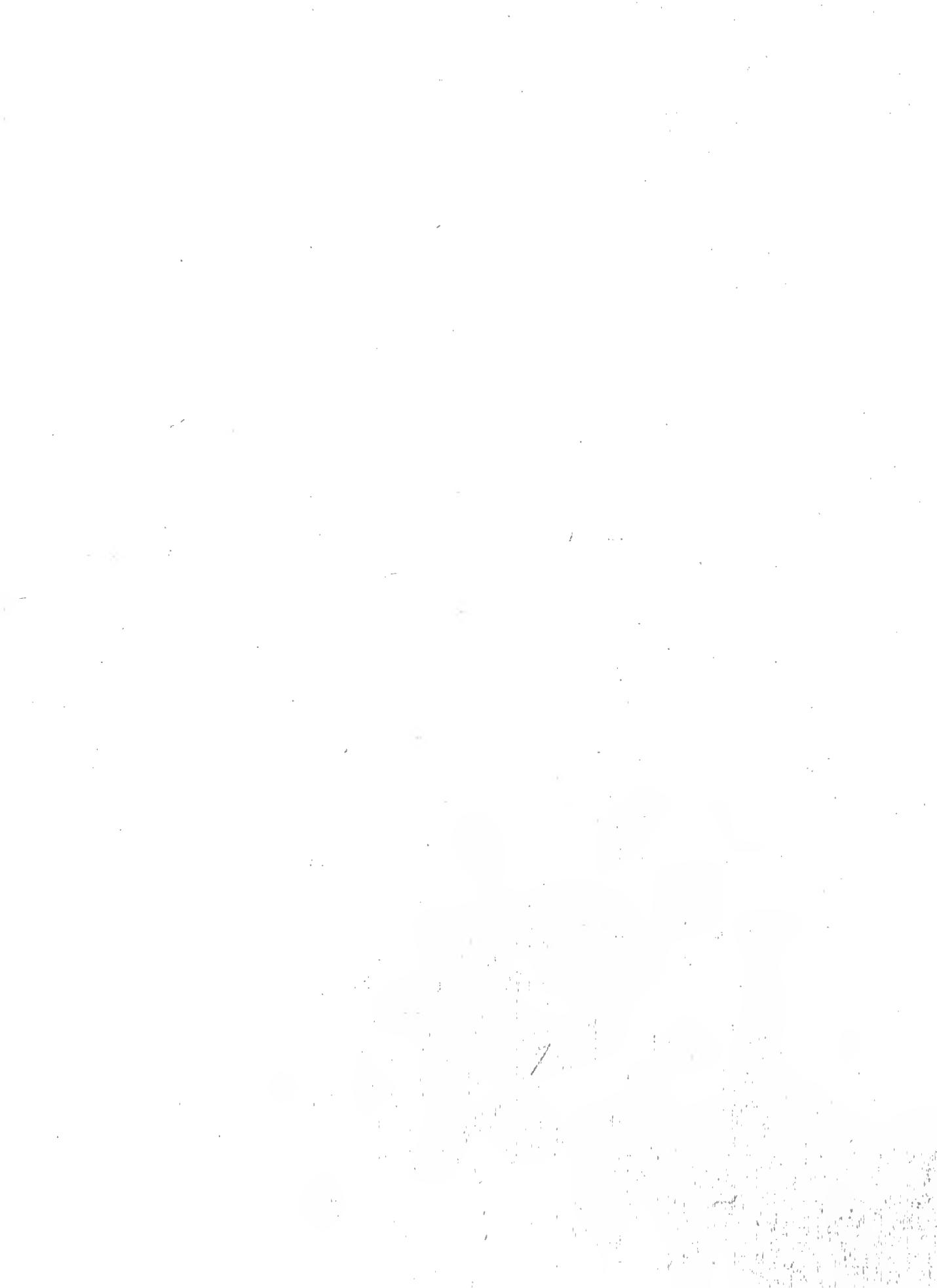
The effects of the magnitude of the ratio of measurement error variance on locating the position of the estimates of parameters should be investigated further; it may also be useful the finite sample properties of the grouping method studied by using Monte Carlo methods for improving the efficiency of the estimators, Durbin's (1954) instrumental variable method or Bartlett's (1949) three group method can be used to replace Wald's two group method. The impact of measurement errors on the functional form should be investigated to understand the interaction effect between the specification bias and the measurement errors. The methods obtained in this paper will be extended to estimate the parameters of a simultaneous equation system in the further research.
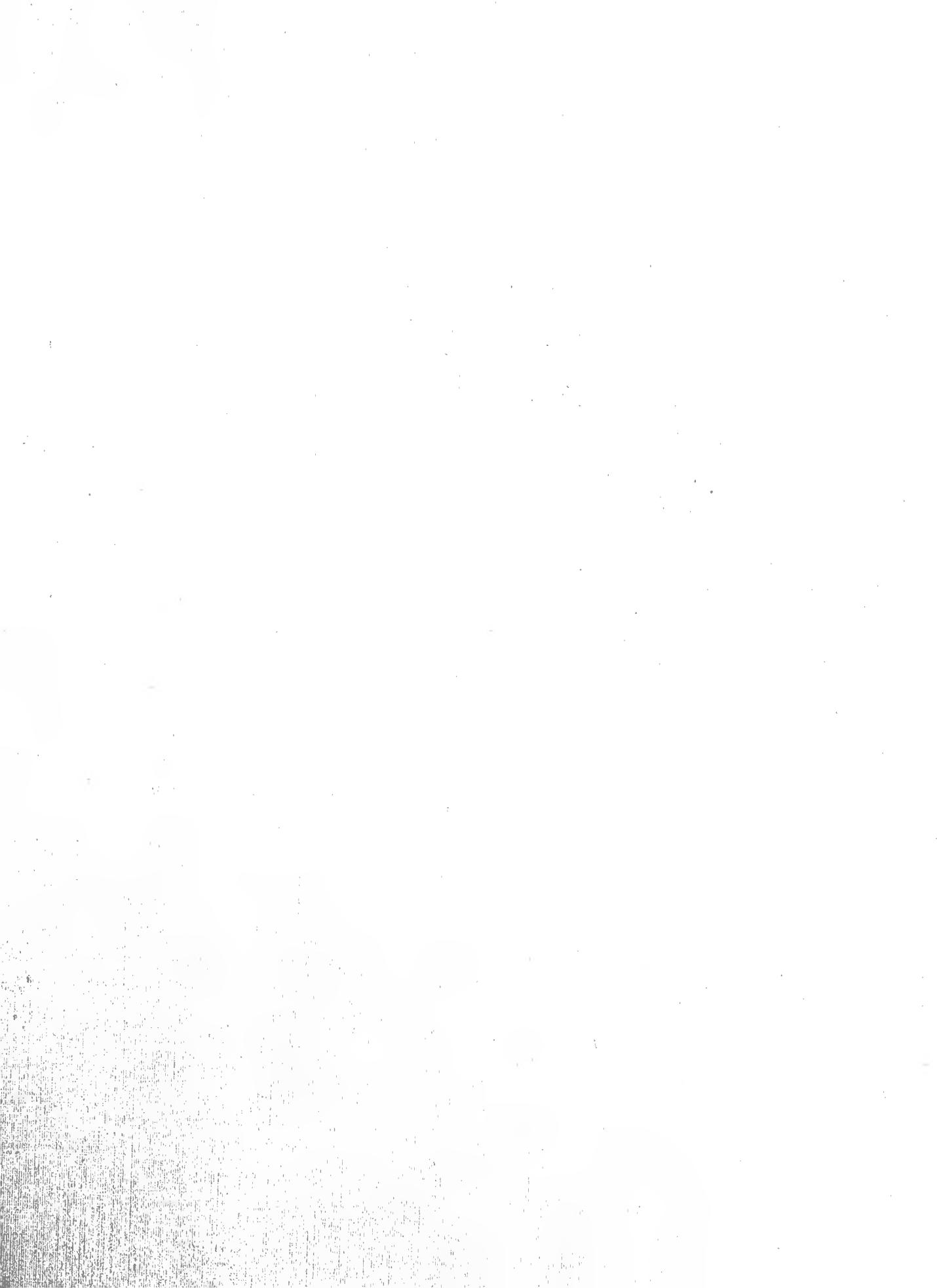
References

1. Allen, R. G. D. (1940), "The Assumptions of Linear Regression," Econometrica, May 1939.

2. Anderson, T. W. and L. Hurwiez (1949), "Errors and Schocks in Economic Relationships," Econometrica, 17, Supplement, 23-25.

3. Bartlett, M. S. (1949), "Fitting a Straight Lines when both Variables are Subject to Error," Biometrics, 5, 207-212.

4. Clutton-Brock, M. (1967), "Likelihood Distributions for Estimating Functions When Both Bariables are Subject to Error," Technometrica, 9, 261-269.

5. Deming, W. E. (1943), Statistical Adjustment of Data, New York: John Wiley and Sons, Inc.

6. Durbin, J. (1954), "Errors in Variables," Review of International Statistical Institute, 22, 23-32.

7. Frisch, R. (1943), Statistical Confluence Analysis by Means of Complete System, Olso, 1934.

8. Gibson, W. M. and G. H. Jowett (1957), "Three Group Gression II: Multiple Regression Analysis," Applied Statistics, 6, 189-197.

9. Goldberger, A. S. (1972), "Structural Equation Methods in the Social Science," Econometrica, 40, 979-999.

10. Graybill, F. A. (1961), An Introduction to Linear Statistical Models, Volume I, New York: McGraw-Hill Book Co., Inc., 1961.

11. Hogg, R. V. and A. T. Graig (1969), Introduction to Mathematical Statistics, New York: The MacMillan Company, 1969.

12. Hooper, J. W. and H. Theil (1958), "The Extension of Wald's Method of Fitting Straight Lines to Multiple Regression," Review of International Statistical Institute, 26, 37-50.

13. Kendall, M. G. and A. Stuart (1961), The Advanced Theory of Statistics, II, London: Griffin.

14. Konijn, H. S. (1962), "Identification and Estimation in a Simultaneous Equation Model with Errors in the Variables," Econometrica, 30, 79-87.

15. Koopman, T. (1934), <u>Linear Regression Analysis of Economic Time Series</u>, Haarlem.

16. McCallum, B. T. (1972), "Relative Asymptotic Bias of Omission and Measurement," <u>Econometrica</u>, 40, 757-758.

17. Moran, P. A. P. (1971), "Estimating Structural and Functional Relationships," <u>Journal of Multivariate Analysis</u>, 1, 232-255.

18. Reiersol, O. (1941), "Confluence Analysis by Means of Lag Moments and other Methods of Confluence Analysis," <u>Econometrica</u>, 9, 1-24.

19. Schultz, H. (1925), "The Statistical Law of Demand," <u>Journal of Political Economy</u>, 33.

20. Theil, H. (1971), <u>Principle of Econometrics</u>, New York: John Wiley and Sons, Inc.

21. Tintner, G. (1952), <u>Econometrics</u>, New York: John Wiley and Sons, Inc.

22. Wald, A. (1940), "The Fitting of Straight Lines if both Variables are Subject to Error," <u>Ann. Math. Stat.</u>, 11, 284-300.

23. Wickens, M. R. (1972), "A Note on the Use of Proxy Variables," <u>Econometrica</u>, 40, 359-361.

24. York, D. (1966), "Least-Squares Fitting of A Straight Line," <u>Canadian Journal of Physics</u>, 44, 1079-1086.

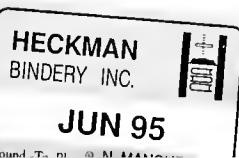25. Zellner, A. (1970), "Estimation of Regression Relationship Containing Unobservable Variables," <u>International Economic Review</u>, 11, 441-454.